Our Docket No.: 42390P10452

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
	)	
Sudheer Sirivara	)	Examiner: Vo, Huyen
	)	
Application No.: 09/822,547	)	**Art Group: 2655**
	)	
Filed: March 30, 2001	)	
	)	
For: Compressing and Using a Concatenative Speech Database In Text-to-Speech Systems	)	

Commissioner of Patents
P.O. Box 1450
Alexandria, VA 22313-1450

DECLARATION UNDER 37 CFR 1.131 IN SUPPORT OF PRIOR INVENTION

Sir :

We declare:

1. Intel Corperation is the assignee of the claims of the above-captioned patent application ("the Application") and of the subject matter described therein.

2. At least prior to January 29, 2001, the filing date of Kochanski et al.U.S. Patent No. 6,625,576 cited in an Office Action mailed May 3, 2005, the invention claimed in the Application had been conceived and reduced to practice in the United States.

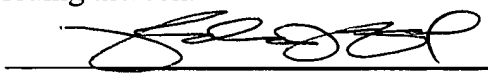
3. The attached Exhibit is an invention disclosure form describing the design of compressing and using a concatenative speech database in text-to-speech systems.

4. The Exhibit discloses that "[i]n order to minimize the size of the concatenative database without significantly affecting the quality of the synthesized speech . . . individual diphone waveforms [were compressed] using a G.723 audio encoder . . ." The Exhibit further discloses that the size of the database was optimized ". . . with respect to the LPC coefficients. The effective compression ratio was approximately 20:1." (Exhibit, page 1, 4-17).

5. Therefore, the Exhibit establishes that the subject matter claimed in the Application had been reduced to practice in the United States prior to January 29, 2001.

Furthermore, all statements made herein of our own knowledge are true and all statements made on information and belief are believed to be true, and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application of any patent issuing thereon.

Dated: July 25, 2005



Aslam A. Jaffery (Attorney for the Assignee)
Reg. No. 51,841

Exhibit

aligned along byte offsets that are recorded & stored while creating the database. Given the set of diphones required to completely map the range of human speech production, the effective size of the concatenative database becomes very large, in the order of 6 MB.

In order to minimize the size of the concatenative database without significantly affecting the quality of the synthesized speech, the following idea was implemented.
The plan was to compress the individual diphone waveforms using a G.723 audio encoder while creating the concatenative database. Since we have access to the diphone waveforms along with the derived residuals, a further optimization in size with respect to the LPC coefficients was achieved. This is because when the G.723 encoder stores the compressed residual it also stores the LPC coefficients in the compressed packet. This optimized the compression of the database, in that we did not need to store the LPC coefficients separately. Care was taken to prevent poor encoding while compressing the diphone by providing guard bands around the diphone. This ensured that the encoder was primed before it started compressing the diphone. In one embodiment the G.723 encoder was set to encode at 5.3 kbps rate, which accepted 240msec frames of audio & compressed them into 20byte packets. The effective compression ratio was approximately 20:1. Fig. 3 illustrates the compression scheme.

Once the database was compressed, the waveform synthesizer needed to be modified to work with the compressed database. Fig. 4 illustrates the decompression procedure.

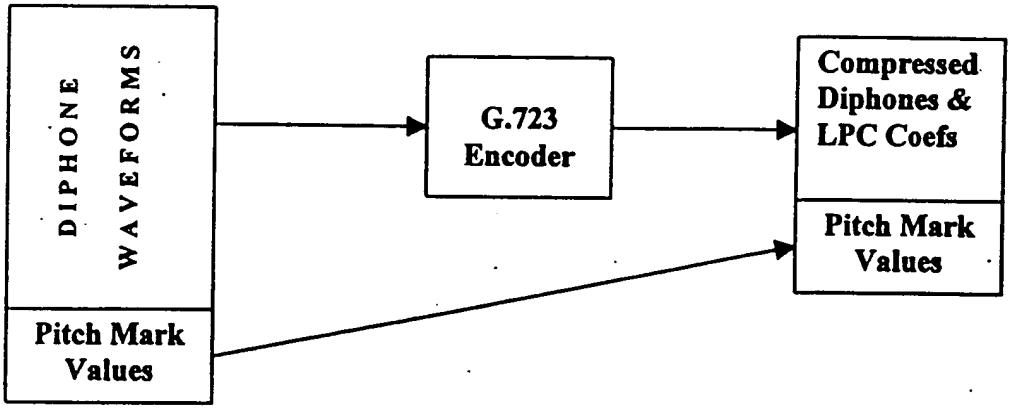


Figure 3: Compression Scheme

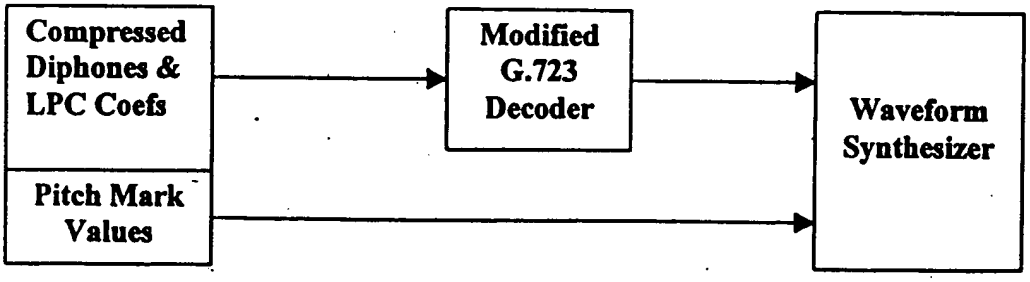


Figure: 4 Decompression Scheme

In order to use the compressed concatenative database, a modified G.723 decoder was employed. The given embodiment of the waveform synthesizer expected diphone residuals to reconstruct the diphone & subsequently concatenate them. Therefore, when

the synthesizer requested a particular diphone, the appropriate compressed diphone was located, based on the offsets recorded during compression & extracted from the compressed packet. However during compression, since the diphone waveforms were compressed, rather than the residuals, the decoder was modified to extract the residual & its LPC coefficients & supplied it to the synthesizer without actually reconstructing the diphone waveform. This ensured that there was no degradation in the quality of the synthesized speech because of the added compression & reconstruction. The pitch mark values, which form a small part of the database, were not compressed & provided directly to the synthesizer.

By employing the compression scheme described above, the size of the concatenative diphone database in one embodiment was reduced from **6.1MB** to about **500kB**. The difference in quality of the synthesized speech with the compressed database, in comparison to synthetic speech from an uncompressed database was minimal.